Generar un archivo tipo Notebook de Python que contenga el código fuente: Uso de librería Pandas para leer un archivo (fifa_eda.csv)

```python
# 1  Generar un archivo tipo Notebook de Python que contenga el código fuente:

#Uso de librería Pandas para leer un archivo (fifa_eda.csv)

import pandas as pd

df = pd.read_csv('fifa_eda.csv')

df.head()
```
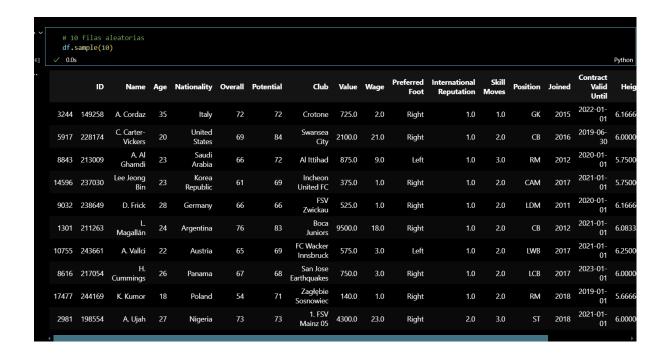✓ 0.0s   Python
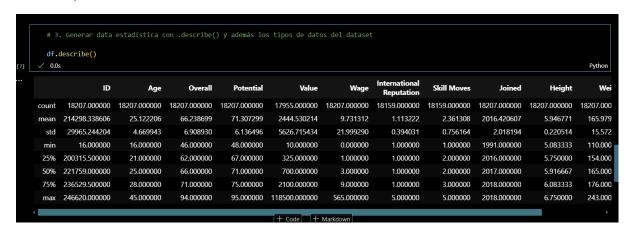
	ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred Foot	International Reputation	Skill Moves	Position	Joined	Contract Valid Until	Height
0	158023	L. Messi	31	Argentina	94	94	FC Barcelona	110500.0	565.0	Left	5.0	4.0	RF	2004	2021-01-01	5.583333
1	20801	Cristiano Ronaldo	33	Portugal	94	94	Juventus	77000.0	405.0	Right	5.0	5.0	ST	2018	2022-01-01	6.166667
2	190871	Neymar Jr	26	Brazil	92	93	Paris Saint-Germain	118500.0	290.0	Right	5.0	5.0	LW	2017	2022-01-01	5.750000
3	193080	De Gea	27	Spain	91	93	Manchester United	72000.0	260.0	Right	4.0	1.0	GK	2011	2020-01-01	6.333333
4	192985	K. De Bruyne	27	Belgium	91	92	Manchester City	102000.0	355.0	Right	4.0	4.0	RCM	2015	2023-01-01	5.916667

Mostrar las primeras 20 filas del archivo, las últimas 5 y un sample de 10

```python
# 2. Mostrar las primeras 20 filas del archivo, las últimas 5 y un sample de 10

df.head(20) #20 primeras filas
```
✓ 0.0s   Python

	ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred Foot	International Reputation	Skill Moves	Position	Joined	Contract Valid Until	Heig
0	158023	L. Messi	31	Argentina	94	94	FC Barcelona	110500.0	565.0	Left	5.0	4.0	RF	2004	2021-01-01	5.5833
1	20801	Cristiano Ronaldo	33	Portugal	94	94	Juventus	77000.0	405.0	Right	5.0	5.0	ST	2018	2022-01-01	6.1666
2	190871	Neymar Jr	26	Brazil	92	93	Paris Saint-Germain	118500.0	290.0	Right	5.0	5.0	LW	2017	2022-01-01	5.7500
3	193080	De Gea	27	Spain	91	93	Manchester United	72000.0	260.0	Right	4.0	1.0	GK	2011	2020-01-01	6.3333
4	192985	K. De Bruyne	27	Belgium	91	92	Manchester City	102000.0	355.0	Right	4.0	4.0	RCM	2015	2023-01-01	5.9166
5	183277	E. Hazard	27	Belgium	91	91	Chelsea	93000.0	340.0	Right	4.0	4.0	LF	2012	2020-01-01	5.6666
6	177003	L. Modrić	32	Croatia	91	91	Real Madrid	67000.0	420.0	Right	4.0	4.0	RCM	2012	2020-01-01	5.6666
7	176580	L. Suárez	31	Uruguay	91	91	FC Barcelona	80000.0	455.0	Right	5.0	3.0	RS	2014	2021-01-01	6.0000
8	155862	Sergio Ramos	32	Spain	91	91	Real Madrid	51000.0	380.0	Right	4.0	3.0	RCB	2005	2020-01-01	6.0000
9	200389	J. Oblak	25	Slovenia	90	93	Atlético Madrid	68000.0	94.0	Right	3.0	1.0	GK	2014	2021-01-01	6.1666
10	188545	R. Lewandowski	29	Poland	90	90	FC Bayern München	77000.0	205.0	Right	4.0	4.0	ST	2014	2021-01-01	6.0000
11	182521	T. Kroos	28	Germany	90	90	Real Madrid	76500.0	355.0	Right	4.0	3.0	LCM	2014	2022-01-01	6.0000
12	182493	D. Godín	32	Uruguay	90	90	Atlético Madrid	44000.0	125.0	Right	3.0	2.0	CB	2010	2019-01-01	6.1666

```python
# 5 ultimas filas

df.tail()
```
✓ 0.0s   Python

	ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred Foot	International Reputation	Skill Moves	Position	Joined	Contract Valid Until	He
18202	238813	J. Lundstram	19	England	47	65	Crewe Alexandra	60.0	1.0	Right	1.0	2.0	CM	2017	2019-01-01	5.75
18203	243165	N. Christoffersson	19	Sweden	47	63	Trelleborgs FF	60.0	1.0	Right	1.0	2.0	ST	2018	2020-01-01	6.25
18204	241638	B. Worman	16	England	47	67	Cambridge United	60.0	1.0	Right	1.0	2.0	ST	2017	2021-01-01	5.66
18205	246268	D. Walker-Rice	17	England	47	66	Tranmere Rovers	60.0	1.0	Right	1.0	2.0	RW	2018	2019-01-01	5.83
18206	246269	G. Nugent	16	England	46	66	Tranmere Rovers	60.0	1.0	Right	1.0	2.0	CM	2018	2019-01-01	5.83

```python
# 10 filas aleatorias
df.sample(10)
```
✓ 0.0s

	ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred Foot	International Reputation	Skill Moves	Position	Joined	Contract Valid Until	Heig
3244	149258	A. Cordaz	35	Italy	72	72	Crotone	725.0	2.0	Right	1.0	1.0	GK	2015	2022-01-01	6.1666
5917	228174	C. Carter-Vickers	20	United States	69	84	Swansea City	2100.0	21.0	Right	1.0	2.0	CB	2016	2019-06-30	6.0000
8843	213009	A. Al Ghamdi	23	Saudi Arabia	66	72	Al Ittihad	875.0	9.0	Left	1.0	3.0	RM	2012	2020-01-01	5.7500
14596	237030	Lee Jeong Bin	23	Korea Republic	61	69	Incheon United FC	375.0	1.0	Right	1.0	2.0	CAM	2017	2021-01-01	5.7500
9032	238649	D. Frick	28	Germany	66	66	FSV Zwickau	525.0	1.0	Right	1.0	2.0	LDM	2011	2020-01-01	6.1666
1301	211263	L. Magallán	24	Argentina	76	83	Boca Juniors	9500.0	18.0	Right	1.0	2.0	CB	2012	2021-01-01	6.0833
10755	243661	A. Vallci	22	Austria	65	69	FC Wacker Innsbruck	575.0	3.0	Left	1.0	2.0	LWB	2017	2021-01-01	6.2500
8616	217054	H. Cummings	26	Panama	67	68	San Jose Earthquakes	750.0	3.0	Right	1.0	2.0	LCB	2017	2023-01-01	6.0000
17477	244169	K. Kumor	18	Poland	54	71	Zagłębie Sosnowiec	140.0	1.0	Right	1.0	2.0	RM	2018	2019-01-01	5.6666
2981	198554	A. Ujah	27	Nigeria	73	73	1. FSV Mainz 05	4300.0	23.0	Right	2.0	3.0	ST	2018	2021-01-01	6.0000

Generar data estadística con .describe() y además los tipos de datos del dataset

```python
# 3. Generar data estadística con .describe() y además los tipos de datos del dataset

df.describe()
```
✓ 0.0s

	ID	Age	Overall	Potential	Value	Wage	International Reputation	Skill Moves	Joined	Height	Wei
count	18207.000000	18207.000000	18207.000000	18207.000000	17955.000000	18207.000000	18159.000000	18159.000000	18207.000000	18207.000000	18207.000
mean	214298.338606	25.122206	66.238699	71.307299	2444.530214	9.731312	1.113222	2.361308	2016.420607	5.946771	165.979
std	29965.244204	4.669943	6.908930	6.136496	5626.715434	21.999290	0.394031	0.756164	2.018194	0.220514	15.572
min	16.000000	16.000000	46.000000	48.000000	10.000000	0.000000	1.000000	1.000000	1991.000000	5.083333	110.000
25%	200315.500000	21.000000	62.000000	67.000000	325.000000	1.000000	1.000000	2.000000	2016.000000	5.750000	154.000
50%	221759.000000	25.000000	66.000000	71.000000	700.000000	3.000000	1.000000	2.000000	2017.000000	5.916667	165.000
75%	236529.500000	28.000000	71.000000	75.000000	2100.000000	9.000000	1.000000	3.000000	2018.000000	6.083333	176.000
max	246620.000000	45.000000	94.000000	95.000000	118500.000000	565.000000	5.000000	5.000000	2018.000000	6.750000	243.000

+ Code   + Markdown

```python
# tipos de datos del Dataframe
df.dtypes
```
[8] ✓ 0.0s

```
...  ID                            int64
     Name                         object
     Age                           int64
     Nationality                  object
     Overall                       int64
     Potential                     int64
     Club                         object
     Value                       float64
     Wage                        float64
     Preferred Foot               object
     International Reputation     float64
     Skill Moves                 float64
     Position                     object
     Joined                        int64
     Contract Valid Until         object
     Height                      float64
     Weight                      float64
     Release Clause              float64
     dtype: object
```

Si es necesario, pasar a numéricas por lo menos 2 columnas que contengan números para incluirlas en .describe(). Aplicar las técnicas aprendidas.

```python
# 4.Si es necesario, pasar a numéricas por lo menos 2 columnas que contengan números para incluirlas en .describe().

#Para este caso no se hace necesario ahcer conversiones, ya que todos los datos numericos en efecto son enteros o float.
```
✓ 0.0s   Python

Añadir una columna "Years Playing", que calcule el año
actual menos la columna "Joined"

```python
# 5. Añadir una columna "Years Playing", que calcule el año actual menos la columna "Joined"

df['Years Playing'] = 2024 - df['Joined']
df.head(3)
```

	ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred Foot	International Reputation	Skill Moves	Position	Joined	Contract Valid Until	Height	Weight	Release Clause	Years Playing
0	158023	L. Messi	31	Argentina	94	94	FC Barcelona	110500.0	565.0	Left	5.0	4.0	RF	2004	2021-01-01	5.583333	159.0	226500.0	20
1	20801	Cristiano Ronaldo	33	Portugal	94	94	Juventus	77000.0	405.0	Right	5.0	5.0	ST	2018	2022-01-01	6.166667	183.0	127100.0	6
2	190871	Neymar Jr	26	Brazil	92	93	Paris Saint-Germain	118500.0	290.0	Right	5.0	5.0	LW	2017	2022-01-01	5.750000	150.0	228100.0	7

Buscar y mostrar a todos los jugadores de Mexico

```python
# 6. Buscar y mostrar a todos los jugadores de Mexico

df[['Name','Club']].loc[df['Nationality'] == "Mexico"]
```
[15]   ✓ 0.0s

	Name	Club
306	A. Guardado	Real Betis
329	H. Lozano	PSV
371	J. Corona	FC Porto
397	C. Vela	Los Angeles FC
406	H. Herrera	FC Porto
...	...	...
17373	R. Gutiérrez	Monarcas Morelia
17614	A. Ocejo	Santos Laguna
18037	C. Landa	Tiburones Rojos de Veracruz
18068	J. García	Santos Laguna
18113	R. Pasquel	Deportivo Toluca

366 rows × 2 columns

Ordenar y mostrar los datos por la columna
ReleaseClause (sueldo)

```
# 7. Ordenar y mostrar los datos por la columna ReleaseClause (sueldo)

df[['Name','Nationality','Club','Release Clause','Years Playing']].sort_values(by='Release Clause', ascending=False)[:10]
```
✓ 0.0s

	Name	Nationality	Club	Release Clause	Years Playing
2	Neymar Jr	Brazil	Paris Saint-Germain	228100.0	7
0	L. Messi	Argentina	FC Barcelona	226500.0	20
4	K. De Bruyne	Belgium	Manchester City	196400.0	9
5	E. Hazard	Belgium	Chelsea	172100.0	12
25	K. Mbappé	France	Paris Saint-Germain	166100.0	6
17	A. Griezmann	France	Atlético Madrid	165800.0	10
7	L. Suárez	Uruguay	FC Barcelona	164000.0	10
16	H. Kane	England	Tottenham Hotspur	160700.0	14
11	T. Kroos	Germany	Real Madrid	156800.0	10
30	Isco	Spain	Real Madrid	156200.0	11

Generar un nuevo dataset que contenga el año (joined) y el número de jugadores (groupby)

```
# 8. Generar un nuevo dataset que contenga el año (joined) y el número de jugadores (groupby)

df_joined = df.groupby('Joined').size().rename('Total Players').reset_index().sort_values(by="Joined", ascending=False)
df_joined
```
✓ 0.0s

	Joined	Total Players
21	2018	6569
20	2017	4307
19	2016	3799
18	2015	1336
17	2014	818
16	2013	458
15	2012	340
14	2011	201
13	2010	131
12	2009	78
11	2008	53
10	2007	38
9	2006	18
8	2005	17
7	2004	12
6	2003	13
5	2002	10
4	2001	2
3	2000	2
2	1999	1
1	1998	3
0	1991	1

❖*Opcional*❖: Generar un gráfico que contenga, por año, el número de jugadores

```python
# 9. ❖Opcional❖: Generar un gráfico que contenga, por año, el número de jugadores

import matplotlib.pyplot as plt
ax = df_joined.plot.bar(x='Joined', y='Total Players', rot=0)
plt.xticks(rotation=45)
plt.title('GRAFICO DE JUGADORES POR AÑO')
plt.show()
```
[31] ✓ 0.1s

```
<Figure size 2000x500 with 0 Axes>
```